Further, the DNA is useful for production of the above described enzyme, or serves as a therapeutic tool, diagnostic tool or research tool for the treatment of diseases caused by the abnormal expression of Gb3/CD77, or it may be useful for the treatment or diagnosis of diseases involved in the action of verotoxins.

[0090]

[Sequence Listing]

SEQUENCE LISTING

<110> Seikagaku Corporation

Koichi Furukawa

<120> α1,4-GALACTOSYLTRANSFERASE AND DNA ENCODING THEREOF

<130> P-7040

<140>

<141> 2000-02-14

<160> 2

<170> PatentIn Ver. 2.0

[0091]

<210> 1

<211> 1975

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (134)..(1192)

<400> 1

aaggtcggct gctgagccag ggcgtgtctc ccggaggcct gtgggctgcc aggatcccca 60

cctctctgca atgggctgcc caggctgacc agccggttcc tgctggaagc tcctggtctg 120

atctggggat acc atg tcc aag ccc ccc gac ctc ctg ctg cgg ctg ctc 169

Met Ser Lys Pro Pro Asp Leu Leu Leu Arg Leu Leu

1 5 10

cgg ggc gcc cca agg cag cgg gtc tgc acc ctg ttc atc atc ggc ttc 217

Arg Gly Ala Pro Arg Gln Arg Val Cys Thr Leu Phe Ile Ile Gly Phe

15 20 25

aag ttc acg ttt ttc gtc tcc atc atg atc tac tgg cac gtt gtg gga 265

Lys Phe Thr Phe Phe Val Ser Ile Met Ile Tyr Trp His Val Val Gly

30 35 40

gag ccc aag gag aaa ggg cag ctc tat aac ctg cca gca gag atc ccc 313

Glu Pro Lys Glu Lys Gly Gln Leu Tyr Asn Leu Pro Ala Glu Ile Pro

45 50 55 60

tgc ccc acc ttg aca ccc ccc acc cca ccc tcc cac ggc ccc act cca 361

Cys Pro Thr Leu Thr Pro Pro Thr Pro Pro Ser His Gly Pro Thr Pro

65 70 75

ggc aac atc ttc ttc ctg gag act tca gac cgg acc aac ccc aac ttc 409

```
Gly Asn Ile Phe Phe Leu Glu Thr Ser Asp Arg Thr Asn Pro Asn Phe
                80                  85                  90
ctg ttc atg tgc tcg gtg gag tcg gcc gcc aga act cac ccc gaa tcc 457
Leu Phe Met Cys Ser Val Glu Ser Ala Ala Arg Thr His Pro Glu Ser
          95                 100                 105
cac gtg ctg gtc ctg atg aaa ggg ctt ccg ggt ggc aac gcc tct ctg 505
His Val Leu Val Leu Met Lys Gly Leu Pro Gly Gly Asn Ala Ser Leu
     110                 115                 120
ccc cgg cac ctg ggc atc tca ctt ctg agc tgc ttc ccg aat gtc cag 553
Pro Arg His Leu Gly Ile Ser Leu Leu Ser Cys Phe Pro Asn Val Gln
125                 130                 135                 140
atg ctc ccg ctg gac ctg cgg gag ctg ttc cgg gac aca ccc ctg gcc 601
Met Leu Pro Leu Asp Leu Arg Glu Leu Phe Arg Asp Thr Pro Leu Ala
                145                 150                 155
gac tgg tac gcg gcc gtg cag ggg cgc tgg gag ccc tac ctg ctg ccc 649
Asp Trp Tyr Ala Ala Val Gln Gly Arg Trp Glu Pro Tyr Leu Leu Pro
            160                 165                 170
gtg ctc tcc gac gcc tcc agg atc gca ctc atg tgg aag ttc ggc ggc 697
Val Leu Ser Asp Ala Ser Arg Ile Ala Leu Met Trp Lys Phe Gly Gly
          175                 180                 185
atc tac ctg gac acg gac ttc att gtt ctc aag aac ctg cgg aac ctg 745
Ile Tyr Leu Asp Thr Asp Phe Ile Val Leu Lys Asn Leu Arg Asn Leu
     190                 195                 200
acc aac gtg ctg ggc acc cag tcc cgc tac gtc ctc aac ggc gcg ttc 793
Thr Asn Val Leu Gly Thr Gln Ser Arg Tyr Val Leu Asn Gly Ala Phe
```

205 210 215 220

ctg gcc ttc gag cgc cgg cac gag ttc atg gcg ctg tgc atg cgg gac 841
Leu Ala Phe Glu Arg Arg His Glu Phe Met Ala Leu Cys Met Arg Asp

225 230 235

ttc gtg gac cac tac aac ggc tgg atc tgg ggt cac cag ggc ccg cag 889
Phe Val Asp His Tyr Asn Gly Trp Ile Trp Gly His Gln Gly Pro Gln

240 245 250

ctg ctc acg cgg gtc ttc aag aag tgg tgt tcc atc cgc agc ctg gcc 937
Leu Leu Thr Arg Val Phe Lys Lys Trp Cys Ser Ile Arg Ser Leu Ala

255 260 265

gag agc cgc gcc tgc cgc ggc gtc acc acc ctg ccc cct gag gcc ttc 985
Glu Ser Arg Ala Cys Arg Gly Val Thr Thr Leu Pro Pro Glu Ala Phe

270 275 280

tac ccc atc ccc tgg cag gac tgg aag aag tac ttt gag gac atc aac 1033
Tyr Pro Ile Pro Trp Gln Asp Trp Lys Lys Tyr Phe Glu Asp Ile Asn

285 290 295 300

ccg gag gag ctg ccg cgg ctg ctc agt gcc acc tat gct gtc cac gtg 1081
Pro Glu Glu Leu Pro Arg Leu Leu Ser Ala Thr Tyr Ala Val His Val

305 310 315

tgg aac aag aag agc cag ggc acg cgg ttc gag gcc acg tcc agg gca 1129
Trp Asn Lys Lys Ser Gln Gly Thr Arg Phe Glu Ala Thr Ser Arg Ala

320 325 330

ctg ctg gcc cag ctg cat gcc cgc tac tgc ccc acg acg cac gag gcc 1177
Leu Leu Ala Gln Leu His Ala Arg Tyr Cys Pro Thr Thr His Glu Ala

335 340 345

atg aaa atg tac ttg tgaggggccc gccaggtcac ctccccaacc tgctcctgat 1232
Met Lys Met Tyr Leu
350

ggggcactgg gccgcccttc ccggggaggc aagattgagg gcccgggaga gggaggcccg 1292
agctgccacc gggcttaggc aggctgttga ggagctgtgg gagcaggccc agtgggaggc 1352
tgtggacacc ccgaggacag tgtcctgtct cgaggcaggg ctgacacatg gtgccatagc 1412
cagcggaggg cgctcagtga gtgccccggg ccttctagac aacaggcagg aaggatgaac 1472
ctcagggcac ccccaggtgg tgcggaaagc caggcagttg ggacagaggt gcccacgagg 1532
gcagaggccg gtgctaaggg gatggggaag aagggacaag attcccagag aggagaggag 1592
gctgttggta ggaaagtggc agggctgggg gagacccagc cccaagggtc cggggcggag 1652
gatgctttgt tcttttctgg ttttggttcc tctttcgcgg ggggtggggg aggtcaacag 1712
ggactgagtg gggcagaggc ccagaagtgc cagcctgggg agccgtttgg gggcagcccc 1772
ttctgcccac cccatccttc ttcctctcca gagatgccag gggggcgtgt atgctctgcc 1832
ccttccctca gacaggggct gggtggggag gctctttagg ctcaggagaa gcattttaaa 1892
gaaaccccca ccctgccgcc cgcattataa acacaggaga ataatcaata gaataaaagt 1952
gaccgactgt caaaaaaaaa aaa 1975

<210> 2
<211> 353
<212> PRT
<213> Homo sapiens

<400> 2
Met Ser Lys Pro Pro Asp Leu Leu Leu Arg Leu Leu Arg Gly Ala Pro
1 5 10 15

```
Arg Gln Arg Val Cys Thr Leu Phe Ile Ile Gly Phe Lys Phe Thr Phe
            20                  25                  30
Phe Val Ser Ile Met Ile Tyr Trp His Val Val Gly Glu Pro Lys Glu
        35                  40                  45
Lys Gly Gln Leu Tyr Asn Leu Pro Ala Glu Ile Pro Cys Pro Thr Leu
    50                  55                  60
Thr Pro Pro Thr Pro Pro Ser His Gly Pro Thr Pro Gly Asn Ile Phe
65                  70                  75                  80
Phe Leu Glu Thr Ser Asp Arg Thr Asn Pro Asn Phe Leu Phe Met Cys
                85                  90                  95
Ser Val Glu Ser Ala Ala Arg Thr His Pro Glu Ser His Val Leu Val
            100                 105                 110
Leu Met Lys Gly Leu Pro Gly Gly Asn Ala Ser Leu Pro Arg His Leu
        115                 120                 125
Gly Ile Ser Leu Leu Ser Cys Phe Pro Asn Val Gln Met Leu Pro Leu
    130                 135                 140
Asp Leu Arg Glu Leu Phe Arg Asp Thr Pro Leu Ala Asp Trp Tyr Ala
145                 150                 155                 160
Ala Val Gln Gly Arg Trp Glu Pro Tyr Leu Leu Pro Val Leu Ser Asp
                165                 170                 175
Ala Ser Arg Ile Ala Leu Met Trp Lys Phe Gly Gly Ile Tyr Leu Asp
            180                 185                 190
Thr Asp Phe Ile Val Leu Lys Asn Leu Arg Asn Leu Thr Asn Val Leu
        195                 200                 205
Gly Thr Gln Ser Arg Tyr Val Leu Asn Gly Ala Phe Leu Ala Phe Glu
```

210 215 220

Arg Arg His Glu Phe Met Ala Leu Cys Met Arg Asp Phe Val Asp His

225 230 235 240

Tyr Asn Gly Trp Ile Trp Gly His Gln Gly Pro Gln Leu Leu Thr Arg

245 250 255

Val Phe Lys Lys Trp Cys Ser Ile Arg Ser Leu Ala Glu Ser Arg Ala

260 265 270

Cys Arg Gly Val Thr Thr Leu Pro Pro Glu Ala Phe Tyr Pro Ile Pro

275 280 285

Trp Gln Asp Trp Lys Lys Tyr Phe Glu Asp Ile Asn Pro Glu Glu Leu

290 295 300

Pro Arg Leu Leu Ser Ala Thr Tyr Ala Val His Val Trp Asn Lys Lys

305 310 315 320

Ser Gln Gly Thr Arg Phe Glu Ala Thr Ser Arg Ala Leu Leu Ala Gln

325 330 335

Leu His Ala Arg Tyr Cys Pro Thr Thr His Glu Ala Met Lys Met Tyr

340 345 350

Leu

[Brief Description of Drawings]

[Fig. 1] It shows flow cytometry indicating the expression of Gb3/CD77 by L cells. The left diagram relates to L cells transfected with pCDM8 while the right diagram to L cells transfected with pVTR1/CDM8. The thick line indicates the result of cells stained with mAb38.13 and FITC-conjugated rabbit anti-rat IgG (secondary antibodies) while the thin line the result of cells stained only with

the secondary antibodies (control).

[Fig. 2] It shows TLC charts of glycolipids extracted from cells transiently transfected with α1,4 Gal-T gene.

A: TLC of glycolipids extracted from L cells transfected with pCDM8 (VC) or pVTR1/CDM8 (TF). RBC represents neutral glycolipids extracted from human B red blood cells.

B: TLC immunostaining of Gb3/CD77 by mAb38.13.

[Fig. 3] It shows the hydropathy plot of a polypeptide of the present invention.

[Fig. 4] It shows the α1,4 Gal-T activity in the extracts of transient transfectants of pVTR1.

A: α1,4 Gal-T activity when LacCer was used as an acceptor.

B: α1,4 Gal-T activity when various acceptors were used. PG represents paragloboside.

[Fig. 5] It shows the result of northern blotting of α1,4 Gal-T gene.

A: the upper columns show the results of hybridization with a ^{32}P-labeled probe derived from pVTR1, while the lower columns show the results of hybridization of the same membranes as in A with a β-actin cDNA probe(control).

B: the expression levels of mRNA of α1,4 Gal-T gene were compared among various human tissues. The ordinate represents the percentage of the expression level of a given tissue with respect to the level of heart after correction with the control.

[Fig. 6] It shows flowcytometry of stable transfectant cells.

The left diagram relates to cells transfected with pSV2neo while the right diagram to cells transfected with pVTR1 and pSV2neo. The thin line indicates the number of cells stained with mAb38.13 and FITC-conjugated rabbit anti-rat IgG (secondary antibodies) while the thick line the number of cells stained only with the secondary antibodies (control).

[Fig. 7] It shows the results of MTT assay of L-neo and L-VTR1. The left graph shows the result of L-neo while the right one the result of L-VTR1.

[Fig. 8] It shows the effect of vero toxins on the cell growth.

[Fig. 9] It shows an electrophoresis indicating the result of DNA fragmentation assay.

[Document Name] Abstract

[Object] To provide α1,4-galactosyltransferase to transfer a galactose residue to C4 position of galactose residue of lactosylceramide or galactosylceramide, and DNA coding for the enzyme.

[Solving Means]

The following polypeptides (A) and (B), and DNAs encoding thereof:

(A) a polypeptide consisting of an amino acid sequence represented by the amino acid Nos. 46-353 in SEQ ID NO: 2; or

(B) a polypeptide which comprises an amino acid sequence including substitution, deletion, insertion or transposition of one or few amino acids in the amino acid sequence of (A) and which has an enzymatic activity to transfer a galactose residue from a galactose donor to C4 position of galactose residue of lactosylceramide or galactosylceramide which serves as an acceptor.

[Drawing Selected] Fig.1